

Titulació	Tipus	Curs
4318299 Visió per Computador / Computer Vision	OB	0

### Professor/a de contacte

Nom: Maria Isabel Vanrell Martorell

Correu electrònic: maria.vanrell@uab.cat

### Equip docent

Ramon Baldrich Caselles

Fernando Luis Vilariño Freire

Dimosthenis Karatzas

Pau Rodriguez Lopez

Carlos Boned Riera

Guillem Arias Bedmar

Luis Gomez Bigorda

### Idiomes dels grups

Podeu consultar aquesta informació al [final](#) del document.

### Prerequisits

Un grau en Enginyeria, Matemàtiques, Física o similar.  
Competències de programació en Python.

### Objectius

Coordinador del mòdul: Dr. Ramon Baldrich Caselles

L'objectiu d'aquest mòdul és introduir les tècniques d'aprenentatge automàtic per resoldre problemes de visió per ordinador. L'aprenentatge automàtic tracta de l'anàlisi automàtica de dades a gran escala. Avui en dia s'ajusta als conceptes bàsics de molts mètodes de visió per ordinador, especialment els relacionats amb el reconeixement visual o la classificació de patrons, on els "patrons" engloben imatges d'objectes mundials, escenes i seqüències de vídeo d'accions humanes, per citar-ne algunes.

Aquest mòdul presenta els fonaments i les tècniques més importants per a la classificació de patrons visuals,

centrats principalment en mètodes supervisats. També es tracten temes relacionats com els descriptors d'imatges i la reducció de dimensionalitat. En la mesura del possible, totes aquestes tècniques es proven i s'avaluen en un projecte pràctic sobre descripció d'escenes a partir d'imatges, en funció de les mètriques i procediments estàndards d'avaluació de rendiment com ara corbes de record de precisió i validació creuada de k.

Els resultats d'aprenentatge són:

- (a) Distingiu els principals tipus de tècniques de ML per a la visió per ordinador: supervisat vs no supervisat, generatiu vs. discriminatori, espai de característiques original i nucli de vectors de característiques.
- (b) Conèixer els punts forts i febles dels diferents mètodes, en part après durant la resolució d'un problema de classificació de patrons reals.
- (c) Ser capaç d'utilitzar implementacions de mètodes existents i construir-les des de zero.

El mòdul aprofundeix en dues aproximacions principals per introduir ML al problema de classificació d'imatges. Utilitzant: a) descripció de la imatge amb descriptors definits manualment, b) descripció de la imatge basat en dades. En el primer cas s'utilitza el Bag of Words, en el segon, l'enfocament de l'aprenentatge profund. El contingut de DL es desenvolupa àmpliament proporcionant tant les bases teòriques de les diferents parts de les modernes acrílteres de Neural Networks, com les bones pràctiques per aplicar-lo en aplicacions reals.

## Resultats d'aprenentatge

1. CA06 (Competència) Aconseguir els objectius d'un projecte de visió fet en equip.
2. KA03 (Coneixement) Identificar quins mètodes d'aprenentatge computacional es poden fer servir en funció de les dades per resoldre un problema de visió.
3. KA10 (Coneixement) Seleccionar els millors procediments d'experimentació que cal dissenyar per a l'aprenentatge computacional des de l'entrenament fins a l'avaluació.
4. KA16 (Coneixement) Reconèixer les dimensions de gènere, mediambientals i ètiques dels sistemes de visió i la seva aplicació.
5. SA03 (Habilitat) Aplicar i avaluar tècniques d'aprenentatge computacional per solucionar un problema particular.
6. SA13 (Habilitat) Calcular l'empremta de carboni de qualsevol experiment en què calgui entrenar una xarxa neuronal profunda.
7. SA14 (Habilitat) Detectar biaixos en conjunts de dades d'aprenentatge que permetin evitar la construcció de sistemes que discriminin socialment.
8. SA17 (Habilitat) Preparar presentacions orals que permetin debatre els resultats del desenvolupament d'un projecte de visió.

## Continguts

1. Introducció a l'aprenentatge computacional
2. Entorn experimental
3. Codificacions: SVM i Random Forest
4. Introducció a les Xarxes Neuronals
5. Introducció al Aprenentatge Profund (Deep Learning, DL)
6. Xarxes Neuronals Convolucionals (CNN)
7. Entrenament: preprocessat de dades, inicialització, optimització del gradient
8. Classificació d'imatges
9. Entendre i visualització de CNNs
10. Mètodes eficients per a l'Aprenentatge Profund

## Activitats formatives i Metodologia

Títol	Hores	ECTS	Resultats d'aprenentatge
Tipus: Dirigides			
Sessions teòriques	20	0,8	CA06, KA03, KA10, KA16, SA03, SA13, SA14, SA17, CA06
Tipus: Supervisades			
Sessions de seguiment de projectes	8	0,32	CA06, KA03, KA10, KA16, SA03, SA13, SA14, SA17, CA06
Tipus: Autònomes			
Treball autònom	113	4,52	CA06, KA03, KA10, KA16, SA03, SA13, SA14, SA17, CA06

Sessions supervisades: *(Algunes d'aquestes sessions podrien ser en línia sincròniques)*

- Sessions magistrals, on els professors explicaran continguts generals dels diferents temes. La majoria seran necessaris per a la resolució de problemes.

Sessions dirigides:

- Sessions de projecte, on es presentaran i discutiran els objectius i problemes dels projectes. Els estudiants hauran d'interactuar amb el coordinador de projecte sobre els problemes surgits i les idees per a resoldre'ls. (Approx. 1 hora/setmana)
- Sessions de presentació, on els estudiants faran una presentació oral sobre como han solucionat el projecte i sobre els resultats obtinguts.
- Sesió d'exàmen, on els estudiants seran avaluats individualment, demostrant l'adquisició dels coneixements desenvolupats i la capacitat de resolució de problemes associats.

Treball autònom:

- Els estudiants hauran d'estudiar i treballar autònomament amb els materials derivats de les classes magistrals i de les sessions de projecte.
- Els estudiants treballaran en grup per a resoldre els problemes plantejats en els projectes amb els següents entregables:
  - Codi
  - Informe
  - Presentació oral

Nota: es reservaran 15 minuts d'una classe, dins del calendari establert pel centre/titulació, per a la complementació per part de l'alumnat de les enquestes d'avaluació de l'actuació del professorat i d'avaluació de l'assignatura/mòdul.

## Avaluació

### Activitats d'avaluació continuada

Títol	Pes	Hores	ECTS	Resultats d'aprenentatge
-------	-----	-------	------	--------------------------

Assistència a sessions	0.05	0,5	0,02	CA06, KA03, KA10, KA16, SA03, SA13, SA14, SA17
Exàmen	0.4	2,5	0,1	CA06, KA03, KA10, KA16, SA03, SA13, SA14, SA17
Projecte	0.55	6	0,24	CA06, KA03, KA10, KA16, SA03, SA13, SA14, SA17

La Nota final d'aquest mòdul es calcula amb la següent fórmula:

Nota final = 0.4 x Examen + 0.55 x Projecte+ 0.05 x Assistència

on,

Exàmen: és la nota obtinguda en l'examen final del mòdul (ha de ser  $\geq 3$ ).

Assistència: és la nota derivada de l'assistència a les sessions (mínim 70%)

Projecte: és la nota que posa el coordinador del projecte d'acord amb els resultats del seguiment setmanal del projecte i de les entregues (ha de ser  $\geq 5$ ). Tot d'acord amb criteris específics com ara:

- Participació en les sessions de discussió i en el treball en equip (avaluació entre-membres)
- Entrega de les parts obligatòries i opcionals del projecte.
- Desenvolupament del codi (estil, comentaris, etc.)
- Informe (justificació de les decisions preses en el desenvolupament del projecte)
- Presentació (Presentació i demostració del projecte desenvolupat)

Només els estudiants que han suspès (Nota final  $< 5.0$ ) poden fer l'examen de recuperació.

## Bibliografia

Articles en revistes:

1. Barber, D. "*Bayesian Reasoning and Machine Learning*". Cambridge University Press, 2012.
2. Yoshua Bengio. "*Learning Deep Architectures for AI*". Foundations and Trends in Machine Learning, Vol. 2, No. 1, 2009.
3. Christopher J. C. Burges. "*Dimension Reduction: A Guided Tour*". Foundations and Trends in Machine Learning, Vol. 2, No. 4, 2009.
4. Christoph H. Lampert. "*Kernel Methods in Computer Vision*". Foundations and Trends in Computer Graphics and Vision, Vol. 4, No. 3, 2008.
5. Tinne Tuytelaars and Krystian Mikolajczyk. "*Local Invariant Feature Detectors: A Survey*". Foundations and Trends in Computer Graphics and Vision, Vol. 3, No. 3, 2007.

Llibres:

1. Ian Goodfellow, Yoshua Bengio and Aaron Courville. "*Deep Learning*". 2016. Cambridge, MA, USA: The MIT Press. ISBN: 978-0262035613
2. Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar, "*Foundations of Machine Learning*" MIT Press, 2012. <http://www.cs.nyu.edu/~mohri/mlbook/>
3. Z.H. Zhou. "*Ensemble Methods: Foundations and Algorithms*". Chapman & Hall/CRC, 2012.

Informes tècnics:

1. Criminisi, A. and Shotton, J. and Konukoglu, E. "*Decision Forests for Classification, Regression, Density Estimation, Manifold Learning and Semi-Supervised Learning*". Technical report MSR-TR-2011-114. Microsoft Research, 2011. [http://research.microsoft.com/pubs/155552/decisionForests\\_MSR\\_TR\\_2011\\_114.pdf](http://research.microsoft.com/pubs/155552/decisionForests_MSR_TR_2011_114.pdf)

## Programari

Entorn de programació en Python amb especial atenció a les llibreries de visió per computador i Keras

### Llista d'idiomes

Nom	Grup	Idioma	Semestre	Torn
(PLABm) Pràctiques de laboratori (màster)	1	Anglès	primer quadrimestre	matí-mixt
(PLABm) Pràctiques de laboratori (màster)	2	Anglès	primer quadrimestre	matí-mixt
(TEm) Teoria (màster)	1	Anglès	primer quadrimestre	matí-mixt

PROVISION